BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W. WASHINGTON, D. C. 20024

B70 12088

SUBJECT: Skylab Data Compression - Description of System and Summary of Error Control Test Results - Case 620

DATE: December 31, 1970

FROM: J. E. Johnson

ABSTRACT

A data compression system is being developed for processing Skylab telemetry data at remote sites of the Manned Space Flight Network. Redundant data on the down links received at the sites will be removed by the remote site computers, and only non-redundant data will be transmitted to the Mission Control Center (MCC) in real-time. If the downlinked data is at least 95% redundant (excluding biomedical data, which will not be compressed) all the information content in the data can be transmitted to MCC in real-time. (A 95% redundancy factor is defined to mean that 95% of the sample values have not changed since their previous sample). If the down link redundancy falls below 95%, adaptive program control will be utilized to prevent overflows or queues on the outgoing lines. Any data missed during this time can be recovered and transmitted post-pass. Analysis of Apollo data has shown that only very rarely during peak activity periods has the down link redundancy fallen below 95%.

It is planned to use three high-speed-data lines from each site, each operating at a rate of 7.2 kbps, to transmit the data to MCC. Uncompressed biomedical data will be transmitted on a fourth line. Consideration is being given to replacing these four lines with a single wide-band-data line having a capacity at least as great as the combined capacity of the four high-speed lines.

The compressed data will be protected by a Hamming forward error correction code. This code will be capable of permitting MCC to correct errors occuring in the same 2400-bit block. Tests of the coding scheme over actual lines operating at 2.4 kbps showed a much higher incidence of uncorrectable errors than had been expected. However, the overall average rate of one uncorrectable block in 1250 was judged to be acceptable.

BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W. WASHINGTON, D. C. 20024

B70 12088

SUBJECT: Skylab Data Compression - Description of System and Summary of Error Control Test Results - Case 620

DATE: December 31, 1970

FROM: J. E. Johnson

MEMORANDUM FOR FILE

Large quantities of operational and experimental data will be generated during Skylab flights, on the order of 4×10^9 bits/day (Ref. 1). Using present Apollo techniques, only a small fraction of this data could be returned to the Mission Control Center (MCC) in real-time or near-real-time. To prevent a tremendous backlog of data at the remote sites and permit more timely and efficient operational and experiment control, a real-time data compression system is being developed for Skylab. This will permit transmission of essentially all telemetry data from remote sites to MCC in real-time, or something very close to it (order of minutes). The system will make use of hardware already existing at remote sites, but will require increases in the number of high-speed data lines linking these sites to MCC, and the bit rates transmitted on these lines. The compression system takes advantage of the high redundancy inherent in the spacecraft-to-ground data streams. It will filter out all redundant data, transmitting to MCC only the vaules of parameters that have changed since they last were sampled, along with appropriate parameter identification and time information. The data will be protected by an error correcting code. This memorandum describes the operational environment in which the system will work, summarizes its basic design and key features, and reviews the results of tests run in July, 1970 to evaluate the effectiveness of the error coding scheme.

Skylab Operational Environment

The data collection network for Skylab will be basically the same that currently exists for Apollo. The stations of the Manned Space Flight Network (MSFN) expected to be operational for Skylab are:

(1) Merritt Island, Fla (MIL)

(2) Bermuda (BDA)

(3) Grand Canary Island (CYI)

(4) Madrid, Spain (MAD)

(5) Ascension Island (ACN)

(6) Carnarvon, Australia (CRO)

(7) Honeysuckle Creek, Australia (HSK)

(8) Guam Island (GWM)

(9) Hawaii (HAW)

(10) Goldstone, Calif. (GDS)

(11) Corpus Christi, Tex. (TEX)

(12) Santiago, Chile (SAN) - this is a proposed station which may instead be located in Argentina, Japan, Alaska, or elsewhere.

(13) USNS Vanguard (VAN) - a tracking ship that will be located in the Atlantic primarily to provide launch phase coverage, and will only remain on station for a few days.

All of these stations have Unified S-Band (USB) capability, required to support the CSM. All of them by the time of Skylab will also have VHF capability, required to support the AM (OWS) and the ATM. Some stations may be required to support other missions during the Skylab activity period (chiefly the deep-space stations of GDS, MAD, and CRO, which must support ALSEP and the lunar subsatellite among other missions). However, Skylab support will be prime. Some stations may be capable of simultaneous dual Skylab and unmanned support.

Characteristics of Skylab telemetry links are summarized in Table 1. It is planned to down link real-time Command Service Module (CSM), Airlock Module (AM), and Apollo Telescope Mount (ATM) data during a normal MSFN contact. Recorded data would be dumped simultaneously with real-time data, but not on every contact. Any one MSFN station might have to support up to three real-time links and up to three recorded and dumped links per contact.

Each MSFN site is connected to MCC by high-speed data (HSD) circuits as far as GSFC, and wide-band data (WBD) circuits from GSFC to MSC. Currently there are two HSD lines for spacecraft telemetry from each site operating at a bit rate of 2.4 kbps each. There is also a line for biomedical telemetry data transmitted in an analog format. Starting with Apollo 15, the biomed data will be converted to a digital format and transmitted on an HSD line operating at a bit rate of 4.8 kbps. Also at that time the other data will be consolidated onto a single HSD line operating at 4.8 kbps. Thus, there will be 9.6 kbps of data transmitted to MCC in real-time over two HSD lines. The WBD lines between GSFC and MSC operate at a bit rate of 50.0 kbps. The added bandwidth of this link is necessary since several sites may be transmitting at the same time.

For Skylab, it is planned to increase the number of HSD lines from each site to four. One of these lines would transmit uncompressed biomed data as planned for Apollo 15 and subsequent. The other three lines would each operate at 7.2 kbps and would transmit compressed telemetry data. A total of 21.6 kbps capacity would then be available for compressed data. Alternatively, consideration is being given to consolidating all of the site's telemetry data onto one WBD circuit from each site. This circuit would then be required to have a capacity of at least 26.4 kbps, more if tracking data is to be multiplexed on it. This latter approach may be more economical, but is probably less reliable. An outage on the WBD circuit would be equivalent to loss of the site, whereas loss of an HSD circuit would permit the site to continue to provide real-time support with reduced capability. Addition of a back-up WBD circuit would probably be uneconomical relative to four HSD circuits.

Data Compression System

In Apollo, a relatively small percentage of the down linked parameters are sampled at a reduced rate at the receiving remote site and transmitted to MCC in real-time. Up to 102.4 kbps of real-time data can be received from the CSM and LM (plus dump data), but allowing for overhead somewhat less than 4.8 kbps of this can be accommodated on the presently existing HSD lines from the site. MSC defines many different HSD formats, pre-mission, and MSFN computer programs are developed around these formats. Each format carries a different emphasis. Key parameters and desired sample rates for each parameter are specified for each format. Formats are selected during the conduct of the mission and their identifications are transmitted, pre-pass, to each site. Format selection is based upon mission phase and whatever the MCC flight controllers decide they are most interested in looking at during the upcoming pass. All other data is logged at the remote site on magnetic tape and ultimately returned to MSC by physical tape transport, which takes days or even weeks for ultimate end-point delivery. (A limited amount of data may be returned by HSD post-pass, upon request by MSC).

It is intended to circumvent this rather cumbersome procedure in Skylab by having the 642B telemetry computer at the MSFN site remove redundancy in the down linked data, and transmit to MCC in real-time all the data that has changed in value since its last on-board sampling. On-board sampling rates for Skylab will vary from 0.2 to 320 samples per second, in most cases parameter values will not change between samples. Apollo experience has shown that less than 5% of the parameters are likely to change in value by a measurable amount between consecutive samples even during high activity periods. If the same trend were to hold for Skylab, a redundancy removal system on the ground could permit real-time transmission of all "information" from the site to MCC over transmission facilities having a bandwidth substantially less than the down linked bandwidth. Alternatively, this can be expressed as a "compression ratio" of greater than 20:1. For Skylab, the real-time down link rates will be 51.2 kbps from the CSM and AM, and 72.0 kbps from the ATM. The total real-time rate will be 174.4 kbps. In addition, there may be dump data. The maximum dump rate would be 126.72 kbps from each of two AM recorders and 72.0 kbps from the ATM recorder, for a total of 325.44 kbps. This dump data would be recorded at the site for post-pass data compression after the real-time data compression cycle was finished.

When a site first acquired the Orbital Assembly and had verified that it was receiving good data, an initial measurement list would be compiled and transmitted to MCC. This list would contain the initial value of every parameter appearing in the down links being processed in a standardized order, so that no parameter identification would be necessary. Following transmittal of this list, the actual compression would start. Each sample of each parameter would be compared against the last value of that parameter transmitted to MCC. Samples that had not changed would be discarded. Those that had changed would be stored in buffers, along with parameter identification. Each buffer will consist of 2400 bits, including overhead and error protection coding, and will be transmitted via HSD or WBD lines at a fixed rate - three buffers per second per line if 7.2 kbps lines are used. Each buffer will contain data from only one vehicle. Unless the data were very active, there would be a considerable amount of fill in most buffers. Overflow will not be allowed to occur (see next paragraph). When allowance is made for overhead, paramater identification, and error control coding, a 21.6 kbps total output rate can support a 174.4 kbps real-time input rate provided that the data is at least 95% redundant (a compression ratio of 20:1 or greater).

If the information content of the data threatened to increase beyond the output capacity, adaptive program control would be used to keep within it. The following steps would be taken as necessary in the order indicated:

1 - The magnitude of change in parameter values (corridor width) considered to be non-redundant would be increased from one PCM count to some larger value. This new corridor width will not be the same for each parameter, and will be varied according to a predefined list. Each vehicle will be assigned a priority, and these corridor width changes would be applied first to the lowest priority vehicle alone, then to successively higher priority vehicles as necessary.

2 - The next step would be to reduce the sampling rate inside the computer on selected parameters with lower priorities. For example, some parameters down linked and input to the computer at 100 samples/sec. might be looked at only 50 times/sec.

3 - If still further reduction were needed, lower priority parameters would be deleted entirely from real-time processing.

The identification of priorities, corridor widths and sampling rates would be supplied by MSC pre-mission for all parameters and vehicles. The more extreme adaptive control measures to avoid overflow would most likely only be required in the event of loss of one or more of the output lines, not from vehicle activity.

The adaptive processing mentioned above would apply only to the output of data for real-time transmission to MCC. Concurrently, an "all data digital tape" will be prepared containing all non-redundant data without regard to the adaptive control procedures needed to prevent real-time overflow. This tape may then be played back to MCC post-pass with the tape speed controlled to match the output line capacity. Dump data would also be handled this way. The dump data would not be processed during the pass, but recorded ahead of the computer. It would be input from the recorders while the all data real-time tape was being transmitted. A second all-data tape would be prepared for the dump data, and this would be transmitted to MCC after the real-time tape transmission was finished. It is estimated that a worst-case data load would require about 50 minutes of post-pass activity at a site. Post-pass processing can be performed using both of the remote site 642B computers;

during the pass one of these computers must be reserved for command processing.

Biomedical data has a very low redundancy factor, so at least initially will not be subjected to compression. The biomedical data (EKG's and respiration waveforms) will be stripped from the down linked data prior to entering the computer, and will be transmitted to MCC on its own dedicated HSD line. A 4.8 kbps rate on this line is adequate to transmit all of the biomed data. Compression algorithms more complex than the one to be used for all the other Skylab data are being studied for possible later use with biomed data.

The error coding to be applied to the output data is termed a Hamming forward error correction code. Each buffer will consist of 2040 data bits (which may include fill) protected by 300 code bits. The remaining 60 bits are for overhead and will not be protected. This code will permit detection and correction of up to 30 errors per buffer. The data will be formatted into 30-bit columns. An error in any column may be corrected if no other error also occurs in that column. All errors caused by noise bursts of up to 30 bits in length could be corrected if no other errors occurred in the same buffer. Ninety of the code bits will be used in a "diagonal parity" check designed to detect cases of two or more errors per column. These cannot be corrected. The coding is designed so that the probability of a buffer with errors escaping detection will be exceedingly low.

If MCC should receive a buffer with uncorrectable errors, it would indicate a request to the transmitting site for a new initial measurement list for the vehicle whose data was in the uncorrectable buffer. When the site received this request, it would stop compression processing on that vehicle's data and generate a new initial measurement list for the vehicle requested from its current data. No attempt would be made to retrieve the data MCC received in error. This means that a transient in the data or an impulse-type event could escape detection in real-time if the parameter subsequently returned to its original value. It would be recorded, however, on the all data digital tape.

<u>Error Coding Tests</u>

GSFC has run several tests to evaluate the effectiveness of the Hamming error correction code described in the previous section to protect compressed data. The most complete series of tests were run July 21-24, 1970 between GSFC and MSFN sites at Guam and Bermuda. Data was transmitted from GSFC to the sites on 2.4 kbps HSD lines, and "looped" back to GSFC. The data was not generated by the Skylab compression program, since that program is not yet available. A 2.4 kbps rate was used because reliable 7.2 kbps operation of the lines and modems was not possible with the facilities available at that time. Correctable errors were corrected and the received data compared against a copy of what had been sent out. The tests lasted about 16 hours total, with about 10.5 hours of complete data recording.

Test results (from Ref. 2) are summarized in Table 2. A buffer containing one or more errors occurred on an average of about once per 200 buffers; one that could not be corrected occurred about once every 1250 buffers. No buffers contained "undetected" errors when compared with the original data available at GSFC. Line bit error rates varied from about 5×10^{-5} to less than 1×10^{-6}, and averaged about 4×10^{-6}. Most of the errors occurred in bursts.

These tests results are compared against the theoretically expected results in Table 3. The theoretical results (from Ref. 3) assume random, single bit errors. The comparison shows this is an unrealistic assumption. For the tests as a whole, the uncorrectable buffer rate was about three orders of magnitude greater than random error theory would predict. Fig. 1 is a plot of the test results in terms of uncorrectable buffer rate vs line error rate. Approximate trend lines have been sketched in (no attempt at a least-squares or other type fit).* It is uncertain as to whether or not enough representative data was obtained to permit developing meaningful conclusions. Furthermore, the extrapolation of results at a 2.4 kbps rate to a 7.2 kbps rate or any other rate is risky. A correctable noise burst of 10^{+} bit-duration at 2.4 kbps will obliterate a buffer transmitted at 7.2 kbps because it will cover more than 30 bits and will exceed the correction capability of the Hamming code. Noise burst statistics need to be examined to make a valid extrapolation of the test results to other transmission rates. All other things being equal, higher bit rates will be more susceptible to errors (impulsive or in bursts) since there will be less transmitted energy per bit at higher rates.

It is not clear at this time how high a buffer rejection rate would be operationally acceptable. The overall rejection rate in these tests averaged about 5.3 buffers/hr. Two buffers/sec. were transmitted. Skylab plans call for transmitting 9 buffers/sec. If the same rejection rate held, about 24 buffers/hr. would be rejected. However, the rate per buffer can be expected to be higher at 7.2 kbps than at 2.4 kbps for the

---

*In drawing the trend line for Bermuda, the results of test #1 were not considered. This test had a particularly long noise burst of about 1000 bit lengths. Leaving out this data point permits drawing a straight trend line; however, it is recognized that this may amount to "throwing away" data.

reasons mentioned above (less noise energy required to cause an error, shorter bursts causing uncorrectable errors). It appears highly desirable to run additional tests at higher rates when facilities permit.

Conclusions

The data compression plan for Skylab appears reasonable in terms of the magnitude of data to be collected and the remote site processing capabilities. The present Apollo-type system of retrieval of data is ineffecient and could be intolerable for Skylab. Two additional high-speed-data lines per site will be required, and a higher transmission bit rate will be needed. Tests of the error correction coding to be used to protect the data disclosed an unexpectedly high incidence of uncorrectable errors. These tests, which were run using a line transmission rate of 2.4 kbps, should be repeated at the anticipated operational rate when facilities permit.

J. E. Johnson

2034-JEJ-ms J. E. Johnson

Attachments

Fig. 1
Tables 1, 2 & 3

REFERENCES

1. Skylab Communications Coverage and Data Handling for the "Pick-a-Day" Study, Bellcomm Memorandum for File, B70 09031, J. E. Johnson, September 4, 1970.

2. Transmission Line Test Using Forward Error Detection/ Correction, UNIVAC memorandum to GSFC, August 18, 1970.

3. Apollo Program Design Report for Binary Detection Corrector, UNIVAC Corp., May 12, 1970 (Revised May 20, 1970).

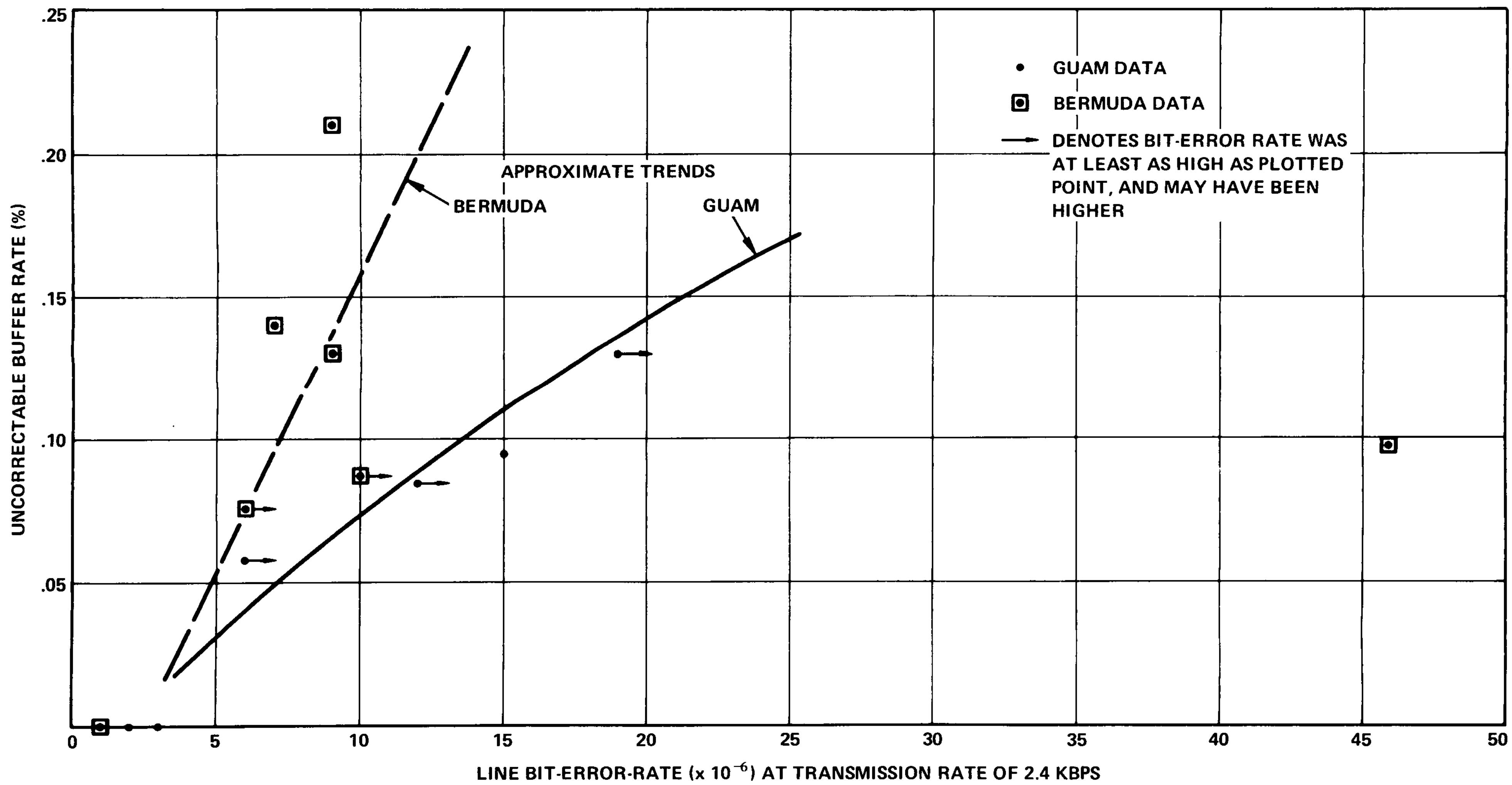

FIGURE 1 - UNCORRECTABLE BUFFER RATE VS LINE BIT-ERROR-RATE
(DATA FROM TESTS RUN JULY 21 – 24, 1970)

Table 1

Skylab Telemetry Down Link Characteristics

Vehicle	Real-Time (RT)/ Dump (D)	Bit Rate (kbps) Record	Bit Rate (kbps) Transmit	Max. Record Time (Hrs.)	VHF/USB	Comments
CSM	RT	--	51.2	--	USB	Either 51.2 or 1.6 KBPS
	RT	--	1.6	--	USB	
CSM-DSE	D	1.6 or 51.2	51.2	0.5 or 2.0	USB	Can be transmitted simultaneously with RT Data
AM	RT	--	51.2	--	VHF	Up to 3 links can be transmitted simultaneously. Normally only one D link will be transmitted simultaneously with RT Data.
	D#1	5.12 or 5.76	112.64 or 126.72	4.0	VHF	
	D#2	5.12 or 5.76	112.64 or 126.72	4.0	VHF	
	D#3	5.12 or 5.76	112.64 or 126.72	4.0	VHF	
ATM	RT	--	72.0	--	VHF	Both links can be transmitted simultaneously.
ATM-ASAP	D	4.0	72.0	1.5	VHF	

DSE-CSM Data Storage Equipment

ASAP-ATM Auxiliary Storage and Playback Equipment

Table 2

Error Correction Code Test Results
July 21-24, 1970
(Data From Ref. 2)

Run		Time (Min.)	Buffers Transmitted	Buffers In Error	Uncorrectable Buffers	Bit Error Rate ($\times 10^{-6}$)	Comments
Guam	1	60	7200	43	9	≥ 19	
	2	70	8508	66	5	≥ 6	
	3	45	5450	9	0	2	
	4	53	6325	18	6	15	Small Bursts
	5	16	2333	5	2	≥ 12	Large Bursts
	6	13	1440	5	0	3	
Bermuda	1	70	8157	45	8	46	Large Bursts, one of 429 bits
	2	60	7017	9	1	7	One large (40 bit) Burst
	3	27	2800	5	0	1	
	4			A B O R T E D			
	5	96	5760	30	5	≥ 10	Large Bursts
	6	41	4800	30	10	9	Small Bursts
	7	39	4760	36	6	9	Very small Bursts
	8	46	5280	50	4	≥ 6	
Totals		636	69910	351	56	Avg $\approx$ 4	
				1:200	1:1250		

Table 3

Comparison of Error Correction Code Results (Experimental vs. Theoretical)

	Theoretical*			Measured
Bit Error Rate	4×10^{-6}	10^{-5}	10^{-4}	4×10^{-6}
Prob. Buffer in Error	9.6×10^{-3}	2.4×10^{-2}	2.4×10^{-1}	9×10^{-3}
Prob. Uncorrectable Buffer	1.6×10^{-6}	10^{-5}	10^{-3}	1.5×10^{-3}
Prob. Undetected Error	2.6×10^{-6}	10^{-14}	10^{-10}	None
Ratio $\left[\frac{\text{Uncorrectable Buffer}}{\text{Buffer in Error}}\right]$	1.7×10^{-4}	4.2×10^{-4}	4.2×10^{-3}	1.6×10^{-1}

*Assumes Random Bit Errors